KYBERNETIKA — VOLUME 11 (1975), NUMBER 5

Fuzzy Metrics and Statistical Metric Spaces

IVAN KRAMOSIL, JIŘÍ MICHÁLEK

The aim of this paper is to use the notion of fuzzy set and other notions derived from this one in order to define, in a natural and intuitively justifiable way, the notion of fuzzy metric. The notion is then compared with that of statistical metric space and both the conceptions are proved to be equivalent in certain sense.

The adjective "fuzzy" seems to be a very popular and very frequent one in the contemporary studies concerning the logical and set-theoretical foundations of mathematics. The main reason of this quick development is, in our opinion, easy to be understood. The surrounding us world is full of uncertainty, the information we obtain from the environment, the notions we use and the data resulting from our observation or measurement are, in general, vague and incorrect. So every formal description of the real world or some of its aspects is, in every case, only an approximation and an idealization of the actual state. The notions like fuzzy sets, fuzzy orderings, fuzzy languages etc. enable to handle and to study the degree of uncertainty mentioned above in a purely mathematic and formal way. A very brief survey of the most interesting results and applications concerning the notion of fuzzy set and the related ones can be found in [1].

The aim of this paper is to apply the concept of fuzziness to the clasical notions of metric and metric spaces and to compare the obtained notions with those resulting from some other, namely probabilistic statistical, generalizations of metric spaces. Our aim is to write this paper on a quite self-explanatory level the references being necessary only for the reader wanting to study these matters in more details.

Definition 1. Let X be a non-empty set. A *fuzzy set* A in X is a pair $\langle X, f_A \rangle$ where f_A is a function defined on X and taking its values in the set $\langle 0, 1 \rangle$ of reals.

Intuitively speaking a fuzzy set is defined, supposing a set X is given and to every element of this set a real non-negative, not greater than 1, is ascribed expressing the degree or the likelihood of the membership of this element in the considered fuzzy set. The sets in the "classical" sense can be considered as a special case of fuzzy sets, especially those for which f_A takes only values 0 or 1, as in this case f_A reduces to a characteristic function defining a subset of the set X. Therefore f_A is usually called the *(generalized) characteristic function* of the fuzzy set A in X.

In the following the basic set X is considered to be fixed, so instead "fuzzy set in X" only the term "fuzzy set" will be used, f_A being an exhaustive characteristic of the fuzzy set A.

Definition 2. The binary relations of equality $(=)$ and inclusion $(\subset)$, unary operation of forming the complement and binary operations of forming the union and intersection $(\cup, \cap)$ for fuzzy sets are defined as follows:

(1) $$A = B, \quad \text{if for all} \quad x \in X \quad f_A(x) = f_B(x),$$

(2) $$f_{A^c}(x) = 1 - f_A(x) \quad \text{for all} \quad x \in X,$$

(3) $$A \subset B, \text{ if } f_A(x) \leqq f_B(x) \quad \text{for all} \quad x \in X,$$

(4) $$f_{A \cup B}(x) = \text{Max}\,(f_A(x), f_B(x)) \quad \text{for all} \quad x \in X,$$

(5) $$f_{A \cap B}(x) = \text{Min}\,(f_A(x), f_B(x)) \quad \text{for all} \quad x \in X.$$

Clearly, these relations and operations are a generalization of those set-theoretic ones and reduce to them supposing f_A, f_B take only the values 0 or 1. As an illustration, the well-known de Morgan laws are valid even for fuzzy sets. Actually,

$$(A \cup B)^c = A^c \cap B^c,$$

as

$$1 - \text{Max}\,(f_A, f_B) = \text{Min}\,(1 - f_A, 1 - f_B)$$

and

$$(A \cap B)^c = A^c \cup B^c$$

as

$$1 - \text{Min}\,(f_A, f_B) = \text{Max}\,(1 - f_A, 1 - f_B).$$

Lemma 1. Let A, B be fuzzy sets. Then $A = B$ if and only if the systems of sets

$$\mathcal{S}_A = \{\{x: x \in X, f_A(x) < \alpha\}, \alpha \in (0, 1\rangle\},$$

$$\mathcal{S}_B = \{\{x: x \in X, f_B(x) < \alpha\}, \alpha \in (0, 1\rangle\}$$

are identical for every $\alpha \in (0, 1\rangle$.

Proof. If $A = B$, then for all $x \in X, f_A(a) = f_B(x)$, so for $\alpha \in (0, 1\rangle$

$$\{x: x \in X, f_A(x) < \alpha\} = \{x: x \in X, f_B(x) < \alpha\},$$

which implies the identity of $\mathscr{S}_A$ and $\mathscr{S}_B$. In the opposite direction equality $\mathscr{S}_A = \mathscr{S}_B$ assures $f_A(x) = f_B(x)$ for all $x \in X$, hence $A = B$. Q. E. D.

This assertion enables to characterize a fuzzy set A up to relation of equality of fuzzy sets by the system $\mathscr{S}_A$ of classical sets.

The notion of fuzzy set can be directly extended in such a way that the notions of fuzzy relation and fuzzy function will be obtained. A "classical" n-ary relation R defined on the set X is clearly defined by a subset of the Cartesian product X^n, namely by the subset of those n-tuples of elements of X, for which relation $R(x_1, x_2, \ldots, x_n)$ holds.

Definition 3. An n-ary fuzzy relation R in the set X is a fuzzy set in the set X^n, i.e. it is a pair $\langle X^n, f_R\rangle$, where f_R maps the n-tuples of elements of X into $\langle 0, 1\rangle$.

Again, considering only the function on X^n, the values of which are only 0 or 1, the notion of fuzzy relation reduces to that of "classical" relation.

Now, it is the very time for us to concern our attention to the notion of metric and metric spaces. We shall start from the well-known definition of usual metric.

Definition 4. Metric ϱ on the set X is a function defined on the Cartesian product $X \times X$ and taking its values in the set E_1 of reals such that the following conditions are valid:

(1) $\varrho(x, y) \geqq 0$ for all $x, y \in X$ (positivity),

(2) $\varrho(x, y) = 0$ if and only if $x = y, x, y \in X$ (identity),

(3) $\varrho(x, y) = \varrho(y, x)$ for all $x, y \in X$ (symmetry),

(4) $\varrho(x, z) \leqq \varrho(x, y) + \varrho(y, z)$ for all $x, y, z \in X$ (triangle inequality).

It is well-known fact that in practice when measuring a distance we are not able, in general, to measure it precisely. This can be explicitly seen from the fact that measuring several times the same distance the results may differ. Usually the average value is taken as an appropriate approximation in such a case. There are at least two approaches enabling to describe and to handle somehow this situation. The first, probabilistic and statistical approach has been developed already for many years; a brief survey can be found in [2]. The other, fuzzy approach, will be explained later, it seems to be, as far as the authors know, an original one.

The probabilistic approach is based on the idea that the distance $d(x, y)$ of two points x, y is an actually existing real number, however, it is, in general, beyond our powers and abilities to obtain its precise value. Our attempts to measure this distance are, from probabilistic point of view, nothing else than random experiment that can be formally described by random variables. If some conditions are satisfied (e.g. if these random experiments are independent and can be repeated potentially infinite many times) we are able, for every positive ε, δ, to obtain a real d_0 such that the sentence "with a probability at least $1 - \varepsilon$ the distance $d(x, y)$ differs from d_0 by less than δ "will be valid. And the sentence of this type is the maximum we are able to obtain, no other information concerning the value $d(x, y)$ is obtainable.

From this intuitive explanation follows that any actually obtained value representing, in some measure, the value $d(x, y)$ is, in fact, a value taken by a random variable and can be, therefore, characterized by its distribution function. So the following definition seems to be quite understandable.

Definition 5. A statistical metric space over the set X is a pair $\langle X, \mathscr{F} \rangle$ where $\mathscr{F}$ is a mapping ascribing to every pair x, y of elements from X a distribution function $\mathscr{F}(x, y)(.)$ (denoted also $F_{xy}(.)$) under the condition that the following is valid:

(1) $F_{xy}(\lambda) = 1$ for all $\lambda > 0$ if and only if $x = y, x, y \in X$,

(2) $F_{xy}(0) = 0$ for all $x, y \in X$,

(3) $F_{xy}(\lambda) = F_{yx}(\lambda)$ for all $x, y \in X, \lambda \in (-\infty, \infty)$,

(4) If $F_{xy}(\lambda) = 1$ and $F_{yz}(\mu) = 1$, then $F_{xz}(\lambda + \mu) = 1$.

Considering the value $F_{xy}(\lambda)$ as the probability that the obtained value of distance is smaller than λ the conditions (1) – (3) of the foregoing definition can be seen to be direct generalizations of positivity, identity and symmetry conditions in the definition of "usual" metric. As far as the triangle inequality is considered the situation is not so simple. Condition (4) expresses the weakest request, namely: if we are sure that $d(x, y)$ is smaller than λ and if we are sure, at the same time, that $d(y, z)$ is smaller than μ, then we can be sure that $d(x, z)$ is smaller than $\lambda + \mu$. Clearly, a demand of such a type must be admitted as its ommiting could lead to a contradiction with the usual definition of metric, which ought to be embeddable in our definition of statistical metric spaces supposing the distributions F_{xy} are those concentrated in one point. For more details concerning this case see [3].

However, from the other side there are many reasons supporting the opinion that this extension of the triangle inequality is too weak to lead to some interesting results. Some discussion concerning this problem can be found in [2]. Probably, the most serious objection concerns the fact, that the triangle inequality generalized in such a way does not bring any limitation for those values of λ, for which $F_{xy}(\lambda) < 1$, being for this case vacuously satisfied. If, e.g., $F_{xy}(\lambda) < 1$ for any $x, y \in X$, $\lambda \in (-\infty, \infty)$ (and this is, in general, the case), then triangle inequality is an empty, tautological, condition. In the paper [3] it is proved that for every statistical metric space there exists a function T_F defined by the following equality:

$$T_F(a, b) = \inf (F_{xy}(\lambda + \mu) : F_{xz}(\lambda) \geqq a, F_{yz}(\mu) \geqq b) .$$

This function T_F is defined on Cartesian product $\langle 0, 1\rangle \times \langle 0, 1\rangle$ and, taking its values in the interval $\langle 0, 1\rangle$, satisfies the following conditions:

$$T_F(a, b) \leqq T_F(c, d) \quad \text{for} \quad a \leqq c, b \leqq d , \tag{1}$$

$$T_F(a, b) = T_F(b, a) , \tag{2}$$

$$T_F(1, 1) = 1 , \tag{3}$$

$$T_F(F_{xz}(\lambda), F_{yz}(\mu)) \leqq F_{xy}(\lambda + \mu) . \tag{4}$$

This way of reasoning gave arise the concept of Menger space.

Definition 6. Menger space is a pair $\langle X, \mathscr{F}\rangle$ satisfying the same conditions as statistical metric spaces the condition (4) being replaced by the following one:

$$F_{xy}(\lambda + \mu) \geqq T(F_{xz}(\lambda), F_{yz}(\mu)) \quad \text{for all} \quad x, y, z \in X , \quad \lambda, \mu \in (-\infty, +\infty) , \tag{4M}$$

where T is a binary real function satisfying for all $a, b, c, d \in \langle 0, 1\rangle$

$$T(a, 1) = a , \quad T(0, 0) = 0 \tag{a}$$

$$T(a, b) \leqq T(c, d) \quad \text{for} \quad a \leqq c, b \leqq d \tag{b}$$

$$T(a, b) = T(b, a) \tag{c}$$

$$T(a, T(b, c)) = T(T(a, b), c) . \tag{d}$$

The fuzzy approach to the notion of distance follows from the idea that the distance between two points is not an actually existing real number which we have to find or to approximate, but that it is a fuzzy notion, i.e. the only way in which the distance in question can be described is to ascribe some values from $\langle 0, 1\rangle$ to various sentences proclaiming something concerning this distance. Namely, in the following we shall limit ourseselves to the assertions claiming the considering distance to be smaller than an a priori given real. As a justification for this limitation can serve the following assertion.

Lemma 2. A metric ϱ on the set X is uniquely determined by the following relation $R_\varrho \subset X \times X \times E_1$: for all $x, y \in X$, $\lambda \in E_1$ relation $R_\varrho(x, y, \lambda)$ is valid if and only if $\varrho(x, y) < \lambda$.

Proof. Let ϱ_1, ϱ_2 be two different metrics on X. Then there exists at least one pair $\langle x, y\rangle \in X \times X$ such that $\varrho_1(x, y) \neq \varrho_2(x, y)$, suppose $\varrho_1(x, y) < \varrho_2(x, y)$. Then

$$\langle x, y, \varrho_2(x, y)\rangle \in R\varrho_1 , \quad \text{but}$$

$$\langle x, y, \varrho_2(x, y)\rangle \notin R\varrho_2 , \quad \text{i.e.} \quad R\varrho_1 \neq R\varrho_2 . \quad \text{Q. E. D.}$$

This assertion leads directly to the following definition.

Definition 7. Fuzzy metric R on the set X is a fuzzy set in the Cartesian product $X \times X \times E_1$ the characteristic function f_R of which satisfies:

(1) $f_R(x, y, \lambda) = 0$ for all $x, y \in X$ and all $\lambda \leqq 0$,

(2) $f_R(x, y, \lambda) = 1$ for $\lambda > 0$ if and only if $x = y$,

(3) $f_R(x, y, \lambda) = f_R(y, x, \lambda)$ for all $x, y \in X$ and all $\lambda \in E_1$,

(4) $f_R(x, z, \lambda + \mu) \geqq S(f_R(x, y, \lambda), f_R(y, z, \mu))$, where S is a measurable binary real function defined on $\langle 0, 1\rangle \times \langle 0, 1\rangle$ taking its values in $\langle 0, 1\rangle$ and such that $S(1, 1) = 1$,

(5) $f_R(x, y, \lambda)$ is for every pair $\langle x, y\rangle \in X \times X$ a left-continuous and non-decreasing function of λ such that $\lim f_R(x, y, \lambda) = 1$, if $\lambda \to \infty$.

All the conditions mentioned in the foregoing Definition 7 seem to have a quite natural interpretation. Conditions (1) – (3) generalize the conditions of identity, non-negativity and symmetry in the usual definition of metric. These conditions express also the fact that the properties of identity, non-negativity and symmetry are generalized, but are not subjected to some fuzziness or uncertainty, only the values of distances are fuzzy notions. Clearly, replacing (2) by

(2′) $$f_R(x, x, \lambda) = 1 \quad \text{for all} \quad x \in X \quad \text{and} \quad \lambda > 0$$

we would obtain the notion of fuzzy pseudo-metric.

Condition (4) expresses probably the most weak form of the triangle inequality saying that the likelihood of the fact that the distance between x and z is smaller than $\lambda + \mu$ is a function of likelihoods of the two particular assertions under the condition that if we are sure that $\varrho(x, y) < \lambda$ and $\varrho(y, z) < \mu$ we can be also sure that $\varrho(x, z) < \lambda + \mu$. Of course, this condition may be subjected to the same criticism as in the case of the statistical metric spaces and it is possible to modify our definition of fuzzy metric supposing S satisfies some more conditions.

Finally, (5) expresses the fact that if we believe, in certain degree, that a distance is beyond a limit, we believe also, in the same or greater degree, that this distance is beyond any larger limit. Written in a slightly precized form: if the real $f_R(x, y, \lambda)$ is understood as a degree of certainty that the distance $\varrho(x, y)$ is smaller than λ, it seems to be quite natural to request that for any $\mu \geqq \lambda$ the inequality $f_R(x, y, \lambda) \leqq \leqq f_R(x, y, \mu)$ should be valid. Hence, for every pair $\langle x, y\rangle \in X \times X$ the function $f_R(x, y, \lambda)$ should be a non-decreasing function of λ. The set of discontinuity points of such a function is at most countable, in every point of this type we define the function f_R to be a left-continuous one. Left continuity is chosen to enable to understand $f_R(x, y, \lambda)$ as the degree of our belief that the distance between x and y is smaller than λ. Clearly, it is also possible to suppose f_R to be right-continuous in λ and to interprete the value $f_R(x, y, \lambda)$ as the degree of our belief that the distance in question is smaller than or equal to λ. The natural assumption of finiteness of any distance justifies the condition concerning the limit value of $f_R(x, y, \lambda)$.

Theorem 1. Any fuzzy metric R defined on X is equivalent to a statistical metric space $\langle X, \mathscr{F}\rangle$ in the sense that for all $x, y \in X$ and for all $\lambda \in (-\infty, \infty)$

$$f_R(x, y, \lambda) = F_{xy}(\lambda) .$$

Proof. Let R be a fuzzy metric on X. The conditions imposed to f_R imply that $f_R(x, y, \lambda)$, considered for any fixed pair $\langle x, y\rangle \in X^2$ as a function of λ, posesses all the properties which a distribution function is to posess. Hence, mapping $\mathscr{F}$ ascribing to $\langle x, y\rangle$ the function $F_{xy}(\lambda) = f_R(x, y, \lambda)$ defines a requested statistical metric space $\langle X, \mathscr{F}\rangle$ on the set X. On the other side, considering a statistical metric space $\langle X, \mathscr{F}\rangle$ and ascribing to every triple $\langle x, y, \lambda\rangle$ from $X \times X \times E_1$ the value $F_{xy}(\lambda) \in \langle 0, 1\rangle$ we obtain, as can be easily seen when looking at the condition (1) – (5) of Definition 6, a fuzzy metric.

Corollary. Any special type of statistical metric spaces resulting from imposing some more conditions on the function T is equivalent, in the same sense as above, to the special type of fuzzy metrics resulting from imposing the same conditions on the function S.

Theorem 2. Let R be a fuzzy metric on the set X satisfying the generalized triangle inequality in the sense of Menger

$$f_R(x, y, \lambda + \mu) \geqq T(f_R(x, z, \lambda), f_R(y, z, \mu)) ,$$

where T is continuous and enables, for every $n \geqq 2$, to define a Lebesgue-Stieltjes measure on the Cartesian product $\langle 0, 1\rangle^n$. Then it is possible to construct a random function $R(\omega, x, y)$ defined, for every $\langle x, y\rangle$, on a probability space $(\Omega, \mathscr{S}, P)$ and such that

$$P(\{\omega : R(\omega, x, y) < \lambda\}) = f_R(x, y, \lambda) ,$$

$$P(\{\omega : R(\omega, x, y) < \lambda + \mu\}) \geqq T(P(\{\omega : R(\omega, x, z) < \lambda\}),$$
$$P(\{\omega : R(\omega, y, z) < \mu\}))$$

for every $x, y, z \in X$, $\lambda, \mu \in E_1$.

Proof. Let $x_1, y_1, x_2, y_2, \ldots, x_n, y_n \in X$. For every such $2n$-tuple we define the function

$$F_{x_1 y_1 \ldots x_n y_n}(\lambda_1, \lambda_2, \ldots, \lambda_n) = T_n(F_{x_1 y_1}(\lambda_1), \ldots, F_{x_n y_n}(\lambda_n))$$

where

$$F_{x_i y_i}(\lambda_i) = f_R(x_i, y_i, \lambda_i), \quad i = 1, 2, \ldots, n,$$

and

$$T_n(a_1, a_2, \ldots, a_n) = T(a_1, T(a_2, T(a_3, \ldots, T(a_{n-1}, a_n)\ldots))).$$

The role of the basic space Ω is played by the space of all real functions defined on $X \times X$. The only thing which rests is to check the Kolmogorov consistency conditions. Q. E. D.

The previous assertions deserve some more comment. It is a well-known fact, see e.g. [2], that it is possible to generalize the usual notion of convergence in such a way that the obtained generalization will be adequate for statistical metric spaces. Moreover, if T is continuous and satisfies the conditons of Menger space, then the covergence in the considered statistical metric space implies the Lèvy's convergence of distribution functions, see e.g. [4]. Theorem 1 and its corollary enable, hence, to "translate" the notion of convergence and other resulting from it topological notions into the language of fuzzy sets, fuzzy relations and fuzzy metrics. The authors feel that there is an intuitive difference between probability and fuzziness, even if both of these notions wish to describe some aspects of uncertainty connected with events and notions in surrounding us world. The further process of introducing fuzziness into the topological spaces theory seems to be a way enabling to achieve explicit results claiming this difference.

There are, as far as the authors know, two papers dealing with fuzzy topological spaces. In [5] the author follows the pattern used in the process of abstract definition of topological spaces and investigates in which measure this pattern can be followed supposing that instead of "usual" sets the fuzzy sets are considered. Another idea is explained in [6], where sets are again the "usual" ones, however, the property of belonging to the closure of a set is subjected to a fuzzification, in other words, the closures of sets are fuzzy sets. It is proved, in [6], that this approach leads to some interesting results which are expressible in the terms of probability theory but which have not been studied or proved when some attempts to apply probability theory in topology were considered.

(Received April 8, 1974.)

[1] H. W. Gottinger: Toward a Fuzzy Reasoning in the Behavioral Sciences. Ekonomicko-matematický obzor *9* (1973), 4, 404—422.

[2] B. Schweizer, A. Sklar. Statistical Metric Spaces. Pacific Journal of Mathematics *10* (1960), 313—334.

[3] E. Nishiura: Constructive Methods in Probabilistic Metric Spaces. Fundamenta Mathematicae *67* (1970), 115—124.

[4] B. Schweizer: On the Uniform Continuity of the Probabilistic Distance. Zeitschrift für Wahrscheinlichkeitstheorie *5* (1966), 357—360.

[5] C. L. Chang: Fuzzy Topological Spaces. Journal of Mathematical Analysis and Applications *24* (1968), 1, 182—190.

[6] J. Michálek: Fuzzy Topologies. Kybernetika *11* (1975), 5, 345—354.

Dr. Ivan Kramosil, CSc., Jiří Michálek, prom. mat.; Ústav teorie informace a automatizace ČSAV (Institute of Information Theory and Automation — Czechoslovak Academy of Sciences), Pod vodárenskou věží 4, 180 76 Praha 8. Czechoslovakia.